Preferential centrality as a multi-regional model for spatial interaction and urban agglomeration

Alexander Hellervik[1], Leonard Nilsson, Claes Andersson

Chalmers University of Technology

Abstract

Understanding how transportation networks affect regional development has been a long-standing challenge for modellers in several disciplines, both in research and practice. Approaches span between light-weight accessibility and centrality models, to data-heavy land use-transport interaction models. Centrality models have been increasingly employed to support spatial planning on the city-scale, where such techniques are attractive due to their low requirements of socio-economic and demographic data, while they also maintain representations of essential features such as accessibility. However, it has been less clear if such approaches can be successfully extended from the urban to the regional scale. In this paper we demonstrate how a recently introduced centrality measure – *preferential centrality* – can be used as a modelling framework on the multi-regional scale, while retaining high intra-urban spatial resolution. Centrality is calculated on a zonal level, with local plot characteristics and network travel times as input. Preferential centrality is calculated similarly to Google PageRank and eigenvector centrality, but with preferential growth as an additional component that represents local agglomeration processes. To examine the explanatory power of this approach, we compare computed centrality with empirical land taxation values, using the southern half of Sweden as a case study area. Using a static accessibility model as benchmark, we find that the preferential model has a higher capacity to reproduce empirical patterns, with regard to geographical correlations as well as for probability distributions. Our findings suggest that preferential centrality analysis can have practical value in urban and regional planning contexts, for example when assessing the geographical distribution of impacts from transport infrastructure investments.

[1] alexander.hellervik@chalmers.se

1. Introduction

When determining the potential for urban activity and land use in a location, the amount and quality of transportation opportunities are essential factors. Change in the transportation system is thereby a driver of urban transformation, such as construction of new buildings, changing patterns of housing and employment, or changing customer bases for services (Hansen, 1959; Alonso, 1964). The effects of such changes are frequently analysed as changes in *accessibility* using a number of distinctly different measures (Handy and Niemeier, 1997). Most accessibility measures, however, rely on manual categorization of the types and locations of the attractions (destinations) that the measured accessibility refers to. This tends to create ambiguity, as well as considerable sensitivity to input data availability, and to methods of data interpretation.

Moreover, in a dynamic setting, attractions themselves frequently change in response to changes in the accessibility-related factors, which are computed using these attractions as input data. For example, new road infrastructure can substantially affect the locations of firms and the local level of employment (Gibbons et al., 2019), and Baum-Snow (2007) has shown how highway expansion has contributed to the transformation of population density in US cities. Accessibility modelling using static attractions can therefore not be expected to capture the full dynamics of long-term urban agglomeration processes. To do so it is necessary to internalize these dynamical feedbacks into the model.

Models that aim to capture feedbacks between accessibility and agglomeration have been developed in the purview of two main traditions. The first is Land Use Transportation Interaction (LUTI) modelling, where highly detailed model systems, consisting of several sub-models, interact in data-heavy computer simulations (Wegener, 2004; Acheampong and Silva, 2015). The other tradition is based on theoretical urban economic modelling, where economies of scale and transport costs are central forces, but spatial details are less elaborated (Alonso, 1964; Krugman, 1996; Fujita et al., 1999; Glaeser, 2008).

In LUTI modelling, the integration of numerous sub-models tends to give rise to complicated internal dependencies, as well as a forbidding need for input data. This is not only costly, but also hampers model portability and scalability since data is unevenly accessible, and since models will tend to be constructed around idiosyncrasies in local data availability. It is often the case that additional data is used to cover up for missing or poorly calibrated internal model mechanisms. Even in situations where all necessary data is available, and model calibration seems successful, the quality of model results can still degenerate, for example when forecasting or scenario analysis requires large changes of parameter values.

Theoretical models in urban and regional economics can, by using a range of assumptions and simplifications, stay within a window of mathematical tractability. In this way they provide important in-principle insight into how different forces interact to shape an urban system. However, because of the lack of spatial details, it is less clear how these models can be used to understand and model a specific real region, for example by providing forecasts to guide a planning process.

The limitations of traditional modelling may imply that simplicity and clarity in modelling must be sacrificed to arrive at empirically useful results. As argued by Barthelemy (2016) this must not necessarily be the case, since there exists substantial potential in using minimal models, inspired by statistical physics of complex systems, that are in agreement with empirical urban patterns, while retaining theoretical transparency. The methodology that we develop in this article is clearly based on such an ambition: a minimal, light-weight model where much of the needed information does not enter as data but is derived from a smaller and more robust set of data that is consistently available across time and space.

Our approach is based on a network centrality measure, which we use to analyse the relative importance of urban locations, using only simple geographical data and transportation networks as input. This approach makes it possible to view the urban system as a network of interacting locations (Barthélemy, 2011; De Montis et al., 2013; Andersson et al., 2006), with a detailed geographical representation, combined with simple activity dynamics. There are several common centrality measures that can be used in urban analytics (e.g. Porta et al., 2006; Sevtsuk and Mekonnen, 2012; Hillier and Hanson, 1989): betweenness centrality, closeness centrality, degree centrality, eigenvector centrality (e.g. Agryzkov et al., 2019) and PageRank centrality (e.g. Jiang, 2006). All these measures have different characteristics and can be expected to capture different aspects of an urban system. However, none of them include mechanisms to represent an urban agglomeration process. Therefore, in this study we focus on the newly introduced *preferential centrality* measure (Hellervik et al., 2019) that includes such a mechanism. Preferential centrality is derived as an extension of eigenvector centrality (Bonacich, 1972), with the incorporation of activity feedback to represent agglomeration.

The main contribution of this study is to extend the scope of the analysis of preferential centrality by Hellervik et al. (2019), from the urban to the regional and sub-national level. Calculated centrality values for different parameter settings are compared to land taxation values that are used as a proxy for urban economic activity. The performance of preferential centrality is also benchmarked against that of eigenvector centrality and of a static accessibility measure. The results develop and lend further support to our hypothesis that transportation infrastructure networks strongly influence the spatial structure of the societal systems that they serve – structure that methods based on the preferential centrality measure can reveal. The upshot is that this can be achieved without using large amounts of detailed socioeconomic data. This, in turn, reduces the problems inherent to LUTI models, as outlined above, and opens the door for new types of light-weight models based on widely available data, such as Open Street Map. Such models could be used in many additional roles and settings and they could be integrated into planning processes on local, regional and national scales.

2. Methods and data

2.1. Preferential centrality modelling

To model urban economic activity using preferential centrality the procedure described in (Hellervik et al., 2019) is followed closely. The main equation,

$$a_j = \left(a_j + \alpha R_j\right) \sum_i \frac{a_i f\left(c_{ij}\right)}{\sum_k (a_k + \alpha R_k) f(c_{ik})},$$

is solved iteratively for every zone j in the system, where a_j is the activity for the zone, α is a parameter, R_j is a static local weight measuring the suitability of the zone (see section 2.3 for details) and $f\left(c_{ij}\right)$ is a deterrence function decreasing with increased generalised cost c_{ij}. The resulting preferential centrality value of each zone is taken to be exactly the activity value of the zone when the model has found a stable solution satisfying (within a numerical threshold) the above equation for all zones.

To facilitate convergence of the iterative solution process, a normalization is performed in every step, by dividing all activity values a_j with the maximum value of a_j. This is a modification of the method used in (Hellervik et al., 2019) where normalization was done with regard to the sum of a_j, which lead to divergence for low values of α.

The generalised cost c_{ij} is calculated for all zone pairs and can in principle include any type of impedance hindering the interaction between zones, but we have in this study chosen to limit the analysis to travel time in the road network. Travel times between zones are calculated using shortest paths with edge travel time as weights. The travel time for each edge is taken to be a simple ratio between metric edge length and speed limit. A constant (5 minutes in this study) is added to all travel time calculations to approximatively account for impedances outside of the represented network, such as walking within zones, parking, etc. For the deterrence function we have used $f(c_{ij}) = c_{ij}^{-\beta}$, where β is a parameter.

In principle, the most realistic deterrence function would be obtained by estimation from empirical interaction data. Such an estimation is outside the scope of this study but could be recommended if the method is to be used in practical applications. Care must then be taken, however, to verify that it is actually economic interaction that is being measured, and not only physical interaction. Simple trip counts could potentially be misleading for such a purpose, since different types of trips could carry different economic value.

The centrality calculation is based on the travel time between all pairs of zones in the model. Each zone pair results in a shortest path calculation, which means that the calculation time can become impractically long for systems with many zones, as in the current study where high spatial detail is required for large study areas. To reduce the computational burden without sacrificing model size, we have chosen to use an approximate travel time calculation for zones that are far apart[2].

2.1. Parameters

There are two key parameters to be varied in the preferential centrality model, α and β. By varying these, the interplay between agglomeration and accessibility can be explored within a large space of possible model outcomes.

The parameter α controls the impact of the static local weights. A high value for α gives relatively less importance to the dynamic activity a_j and more importance

[2] The approximation is designed so that errors are proportional to distance between zones, which makes it possible to keep relative error within a bounded range. Using a random sampling of zone pairs from the model setups used in this study, the relative errors were estimated by comparing exact shortest path travel time calculations with approximations. The average relative errors were found to be below 6 %, which can be considered acceptable for the purposes of this study.

to the static local weight. With a very high value of α, a zone's activity will mainly be determined by its accessibility to other zones with high local weights. With a low value of α, the model becomes more preferential, which means that a zone's activity is highly influenced by its accessibility vis-à-vis other zones with high activity. This can be interpreted as the α-value controlling the strength of the activity agglomeration in the model, with a low value corresponding to high agglomeration and a stronger self-reinforcing process, i.e. positive feedback. A decrease in the value of α could for example represent increased levels of refinement in production and services, and to a decreasing fraction of the economy directly tied to natural resources. It may be surmised that α has decreased historically. For a theoretical discussion of possible micro-mechanisms for agglomeration processes, see e.g. Duranton and Puga (2004).

The parameter β is used to control the distance decay of interactions, with increasing values representing a stronger distance decay. Increasing values of β, for example due to lower transportation efficiency, corresponds to a relative increase of short-range interactions. On the other hand, low β-values can be used to model systems, with very efficient technologies for transportation and communication, wherein a high proportion of interactions take place over long distance.

2.2. Benchmark models

For comparison we use two benchmark models that do not involve the agglomeration mechanism, but in all other aspects use the same input data. The first is denoted as the *Accessibility model,* and it uses the deterrence function to calculate activity as a function of accessible and buildable land (static local weights), according to

$$a_j = R_j \sum_i R_i f(c_{ij}).$$

The other benchmark model is the *Eigenvector model,*

$$a_j = R_j \sum_i \frac{a_i f(c_{ij})}{\sum_k R_k f(c_{ik})},$$

which can be solved with iteration in a similar fashion as the preferential model. The resulting activity will correspond to the eigenvector centrality of the dense network of interaction between locations, determined by the deterrence function and static local weights.

2.3. Input data

The modelling approach is based on three main data sources. Road network data, property polygons and property building restrictions.

Roads are imported with preserved topology and attributes from Open Street Map (OSM), using OSMnx (Boeing, 2017). In cases where speed limit data is missing in the source data, default values are used for the particular road type. Non-drivable roads are removed based on road types and attributes. Ferry connections where motor vehicles are allowed are treated as roads with limited speed and a fixed waiting time.

Property polygons are obtained from the Swedish National Land Survey (Lantmäteriet). Properties are the smallest administrative building blocks in the official land registry. The precision and quality of this data is high and consistent, since the purpose is to establish and prove ownership (which needs to be precise and just). This is the first reason for our choice of using property polygons as the spatial entities in this modelling framework. The second reason is that this spatial representation gives us an exact physical correspondence between model zones and empirical zones with data on land taxation. A third reason is that property polygons represent the most fine-grained geographical subdivision in many actual planning contexts.

Property building restrictions for every zone is obtained by retrieving the taxation type code for the property. The type codes are in most cases the effect of local planning restrictions, which means that they can be interpreted both as a factual description of current as well as potential land use.

Using the type codes, some initial filtering is applied to include only relevant property polygons in the model: properties with one- and two-dwelling buildings, properties with multi-dwelling and commercial buildings, and properties with industrial buildings. This means that properties with attributes indicating a lack of buildings are excluded, as well as properties mainly used for agriculture and forestry. Properties designated for special use; such as parks, government buildings, health care and sports facilities, are also excluded.

Taxation units are usually uniquely tied to property polygons, but in certain special cases one taxation unit can refer to several polygons. In these cases, all relevant data is aggregated to a merged property zone. A property is in the normal case made up by a single polygon, but in certain situations it can comprise several polygons. In those cases, data is split proportionally (according to area) between the polygons.

To arrive at zones to be used in the model, properties are dissolved along common borders. This yields modelling zones approximately on the block level (see Figure 1), with an average number of *3.3* properties included. Finally, model zones are connected from the centroid to the nearest driveable road (excluding highways, bridges and tunnels) by a virtual road connection.

A static local weight (R_j) is computed for each zone based on the accessible, permissible and buildable land area, according to

$$R_j = A_j P_j B_j,$$

where A_j is the physical zone area, P_j is the fraction of the area with type codes indicating that building development is allowed, and B_j is the fraction of the area that is accessible for development with regard to the current road network. A zone's maximum value for R_j equals the physical zone area.

B_j is determined by calculating a *30 m* wide buffer along all roads assumed suitable for nearby building development. Suitable roads are defined to exclude motorways, bridges, tunnels and any road with a speed limit above *60 km/h*. The estimated unbuildable extent of all roads are also geometrically subtracted from the developable buffers.

2.4. Land taxation data for validation

As a proxy for urban economic activity, property taxation values are used. Taxation data was obtained from the Swedish National Land Survey (Lantmäteriet), for the taxation year 2018.

Land taxation data has a range of features making it our choice for validation data for urban areas. Compared to population or building density, land value can give a much more distinct picture of urban agglomeration, since different local economic activity levels are clearly related to the willingness to pay for renting or buying real estate. For example, two different urban areas could have the same population or building density but very different economic output. These differences should show up clearly in well-estimated land valuation. Another feature is the absence of an upper bound on the land value. Physical characteristics such as night lights, population or building heights are not able to respond in the same way, or the same pace, as land value, to the underlying economic activity that is really taking place.

The actual market price for every piece of land is not possible to obtain, both because most properties are not sold very often, and because land is in most cases sold together with buildings. This means that any complete set of land value data must be based on an assessment based on sales prices and rents in the area, together with data on the characteristics of the sold, rented and assessed properties. It is not within the scope of this study to perform such an assessment and that is why we instead rely on the official assessment made by the tax agency.

In Sweden, all properties have a determined taxation value (updated every third year) that according to law is required to correspond to 75% of the current total market value (land and buildings) for the property. Partial values are also determined for land and buildings separately. The tax agency has established detailed procedures for how to accomplish these assessments of market value, and when such data is available, the taxation values are based on representative sales prices and rental values.

2.5. Aggregation, filtering and regression

To compare resulting centrality values (a_j) with empirical land taxation data (y_j), we estimate a simple logged regression model,

$$\log y_j^* = b_0 + b_1 \log a_j^* + \epsilon_j \,,$$

where y_j^* and a_j^* are values aggregated to administrative taxation assessment areas (TAA). TAAs are used by the tax authority as a smallest subdivision when estimating key parameters determining local market value. In Figure 1 the typical size of these areas can be compared to model zones and property polygons. The main reason for this aggregation is the strong spatial autocorrelation of land taxation values within the administrative areas.

The centrality models we investigate are designed to represent urban processes and urban areas. Thus, we cannot expect relevant results in rural settings, where other processes dominate how land is used and valued. Therefore, only the zones with a land taxation value above a pre-defined threshold, were included in the aggregation to TAA. The threshold was chosen to be the geometric average for land taxation per area unit for Sweden, since this can be expected to give a good estimate for the scale for the underlying fat-tailed distribution (Nelson et al., 2019).

The filtering and aggregation process results in some TAAs that comprise only a few properties. Since these observations tend to introduce noise and distort the overall picture we apply a filter based on the number of properties per TAA. The choice of threshold in this filter is somewhat arbitrary so we take as a starting point the official Swedish definition of a *locality,* which requires a cluster to have

200 inhabitants (Statistics Sweden, 2016). Since population data is not used in the current study, we translate this to filtering out all TAAs with less than 66 properties, assuming an average of 3 inhabitants per property. Also, we have chosen to exclude very large (> $2x10^7$ m^2) TAAs, because these can only be found in rural contexts.

Access to the sea is a highly valued attribute for a property. In principle this could be handled by introducing additional factors in the regression model. However, the effect of property taxation of access to the sea is very varied which might distort the analysis of centrality. For the purposes of this study, we have instead chosen to remove those TAAs where more than 50% of the properties are within 500 m of the coastline.

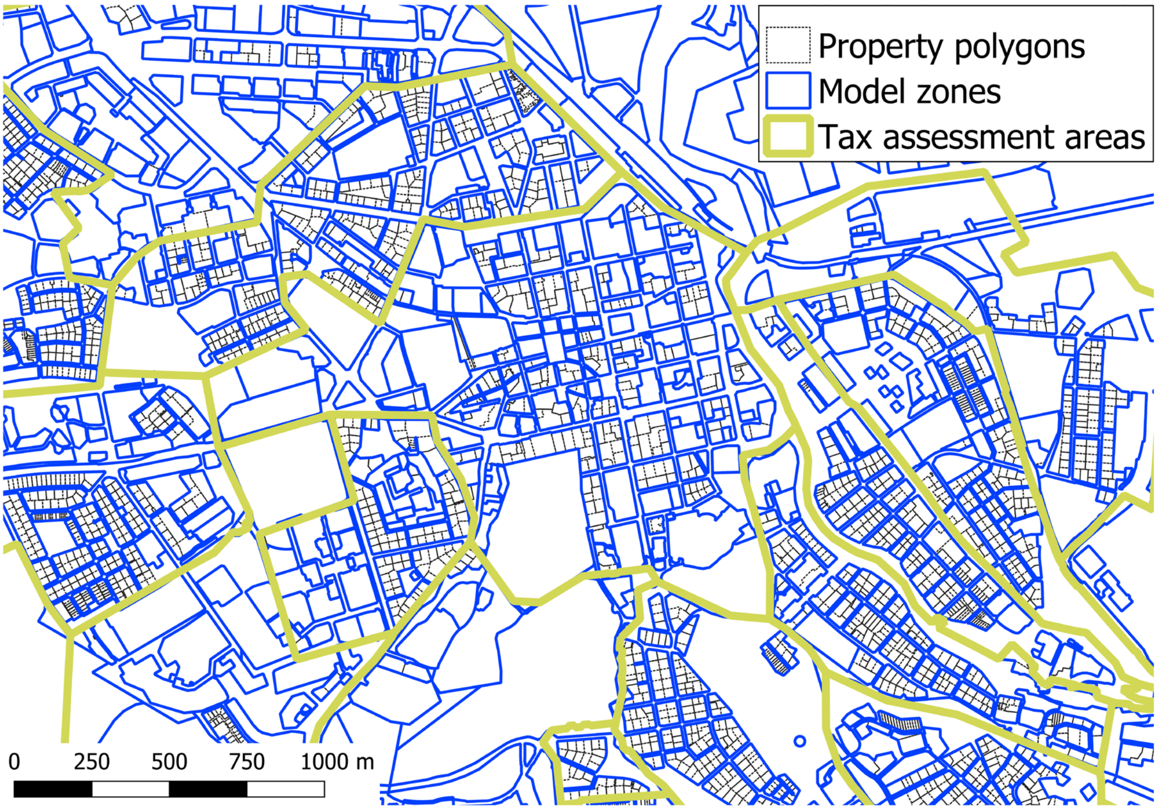

Figure 1 Three levels of spatial detail – property polygons, model zones, and tax assessment areas (TAA). Model zones without visible subdivision contain only a single property polygon.

2.6. Choice of study areas.

We present results both for the single Swedish region Östergötland as well as for a multi-regional study area comprising southern half of Sweden. The northern part of the country is relatively sparsely populated and was excluded to reduce spatial heterogeneity and simplify analysis.

The region of Östergötland was chosen as a suitable study area, since it contains at least two middle-sized cities but no major cities and no significant coastal areas. This is to avoid "monocentric bias" in the analysis, and to demonstrate that the model can handle multiple regional centres.

The chosen study areas are shown in Figure 2, together with a more detailed picture of the geographical patterns of land taxation values, as well an example of corresponding model results.

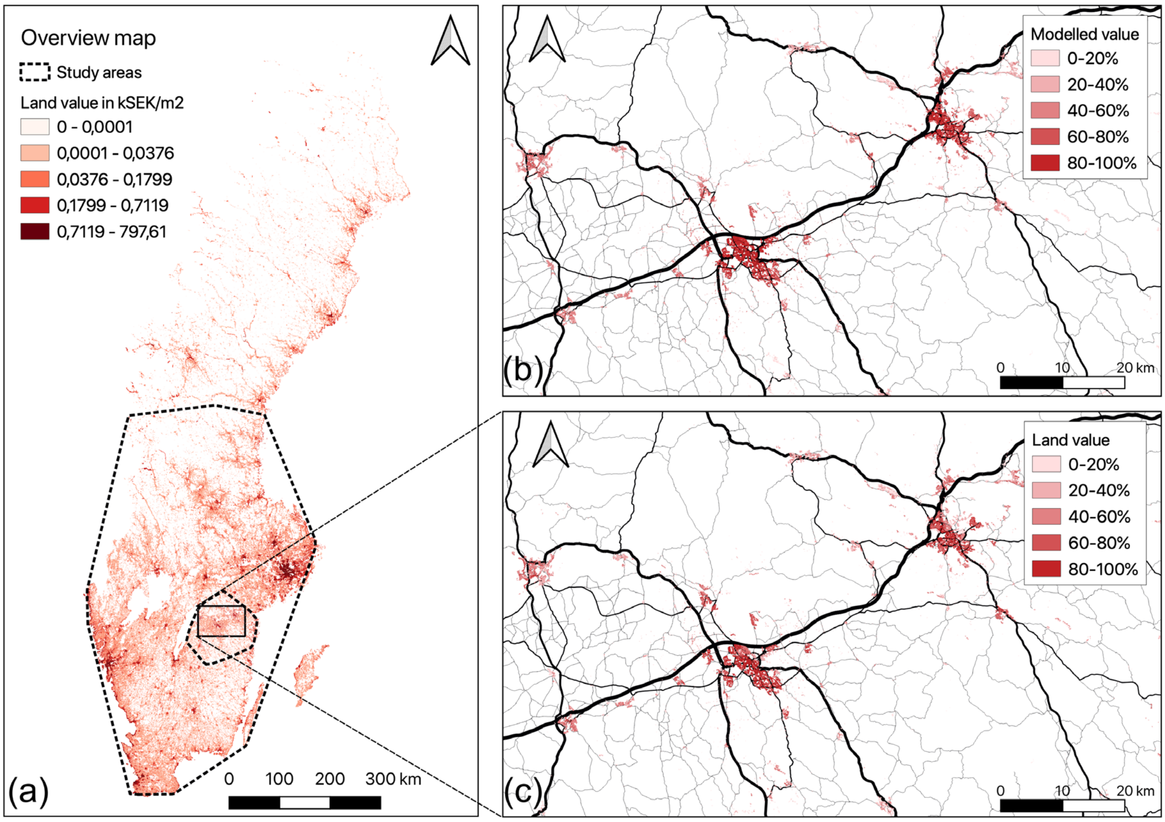

Figure 2. Left panel (a): Land taxation in Sweden, showing the boundaries of the two study areas. Top right panel (b): quantile classification of model results on a zonal level (single region preferential model with $\beta = 2.5$ and $\alpha = 0.06$). Bottom right panel (c): quantile classification of land taxation values for the corresponding model zones.

3. Results

3.1. Distribution of land value

From the complementary cumulative distribution function (CCDF) diagrams shown in Figure 3, it is clear that the preferential model better captures the fat-tailed aspects of the empirical distribution, compared to the eigenvector and accessibility models. The agglomeration parameter α makes it possible to get a distribution of model values in reasonable agreement with empirics.

We have not quantified how the probability distributions fit to the empirical data with any higher degree of exactness since the purpose of this study is not to argue for a certain "true" set of parameters. However, a visual comparison of model and empirical probability distributions alerts us to the importance of studying the tails of the distributions. The benchmark accessibility and eigenvector models cannot produce the heavy tails that characterize the empirical data. The preferential model, however, shows such tails, and Figure 3 serves to illustrate that the general empirical shape of the system coincides with the range of predictions produced by the preferential centrality model.

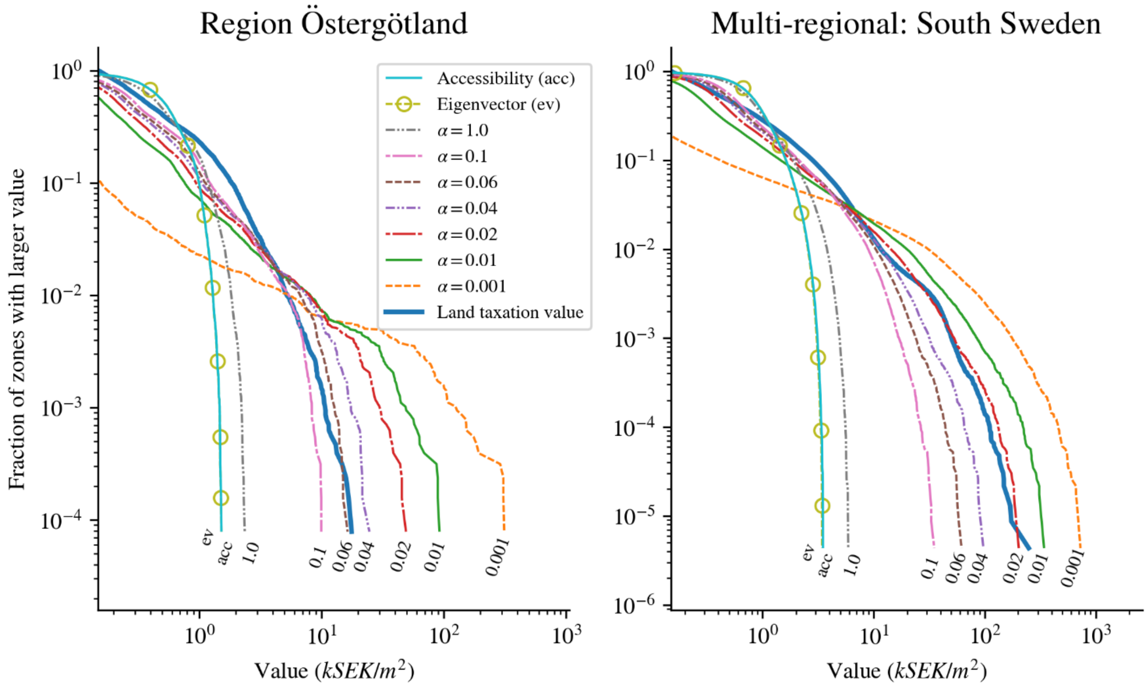

Figure 3. Complementary cumulative distribution functions (CCDF) for: empirical land taxation values (bold line), the preferential model, and the benchmark models – Accessibility and Eigenvector. Note that the lines for Accessibility and Eigenvector coincide. These model results were produced using the interaction parameter setting $\beta = 2.0$. *All zones within 500 m from the coastline were excluded from this analysis.*

3.1. Linear regression

Table 1A compares, for a single geographical region, the performance of the preferential model with different parameter settings, with the eigenvector model and the benchmark accessibility model. Clearly, not all sets of parameter values perform equally well, with performance interpreted as high association (as measured by R^2) between model outcome and empirical land taxation values. The worst performing ($R^2 = 0.11$) combination of parameter values is the preferential model with the highest agglomeration ($\alpha = 0.001$) and the strongest distance decay ($\beta = 3.5$). Low performance is also found on the opposite side of the parametric space for the preferential model with $R^2 = 0.32$ for $\alpha = 8.0$ (low agglomeration) and $\beta = 0.5$ (weak distance decay). These poorly performing parameter values can be interpreted as poorly representing the underlying processes of the empirically observed system. The best R^2-values are found for parameters chosen in the intermediate range, where $\alpha = 0.06$ and $\beta = 2.5$ yield $R^2 = 0.65$.

For all β–values, the simpler models, *Accessibility* and *Eigenvector*, have almost identical performance to each other, and for every column there is a choice of α for which the preferential model outperforms these models. The crucial difference between the simpler models and the preferential model is the presence of an agglomeration mechanism in the latter. This interpretation is reinforced by observing the similarity between results for the highest studied value of α (0.8) with results for the simpler models. With a high α-value the agglomeration process in the preferential model is substantially tuned down, forcing it to behave very similarly to the Eigenvector model (Hellervik et al., 2019).

Table 1B shows the same set of regression results for the multi-regional study area, with similar but somewhat less clear-cut observations, compared to the single region. First, we get a better fit (with regard to R^2) between model and empirical data for the single region. Second, more long-range interaction ($\beta \approx 2{,}0$) gives a better fit for the multi-regional area, compared to the single region where less long-range interaction ($\beta \approx 2{,}5$) yields better performance. This is probably due to boundary effects, causing the distance decay of interaction to affect a smaller study area differently compared to a larger one. A low distance decay cannot play out fully in a small region because the lack of long-range options for interaction.

Further information about the regression results, including estimated parameters and spatial diagnostics, can be found in the Supplementary material. Regression analysis and spatial diagnostics were performed using PySAL (Rey and Anselin, 2007).

3.1. Discussion of results

The presented results provide insights about to what degree spatial patterns in urban activity, here by proxy of land taxation data, may be predicted as patterns of centrality computed using the proposed preferential model. By pattern we mean, more specifically, a combined pattern of frequencies (probability distribution) and locations of urban activity (spatial distribution).

Since the preferential centrality model incorporates agglomeration and interaction as its two principal forces, it is tempting to posit that agglomeration may correspond to frequencies (how much), while interaction would correspond to location (where). The results, however, support a picture where the effects of these forces cannot be easily disentangled. For example, the R^2-values clearly show that both interaction (β) and agglomeration (α) influence the degree of correspondence with the empirical spatial distribution.

The results suggest that it is valuable to include the agglomeration process to improve the performance of centrality models. Preferential centrality modelling makes it possible to capture the basic feedback loop between land use and transportation, while retaining a high degree of simplicity.

Interaction and agglomeration are not separable in the preferential centrality model, and this is probably also true of the real world feedback loops that we aim to represent in the modelling. For example, agglomeration might cause the central district of some specific city to grow. Interaction can then give a spatial spill-over to nearby cities, causing them to grow, which, in turn, is enhanced further by local agglomeration. The end result is a larger amount of regional growth than would have been predicted if interaction and agglomeration were treated separately.

Table 1. R^2-values for a collection of linear regressions comparing model activity values to empirical land taxation. Each row corresponds to a specific α-value (only for the preferential centrality model). The columns represent variation of the parameter β in the deterrence function. For Table 1A, the regressions are based on 185 tax assessments areas and for Table 1B the number is 2831. High values for R^2 are indicative of a better correspondence between model and empirics.

A) Single region: Östergötland

Model version	α	β = 3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	0,11	0,37	0,54	0,43	0,40	0,38	0,37
"	0.002	0,16	0,41	0,56	0,46	0,43	0,40	0,39
"	0.004	0,23	0,47	0,59	0,49	0,46	0,43	0,41
"	0.006	0,29	0,51	0,60	0,51	0,48	0,45	0,43
"	0.008	0,33	0,54	0,61	0,53	0,49	0,46	0,44
"	0.01	0,35	0,55	0,61	0,54	0,50	0,47	0,45
"	0.02	0,43	0,59	0,63	0,57	0,52	0,49	0,48
"	0.04	0,46	0,60	0,64	0,61	0,55	0,52	0,52
"	0.06	0,47	0,61	0,65	0,63	0,58	0,54	0,55
"	0.08	0,48	0,61	0,65	0,64	0,60	0,56	0,56
"	0.1	0,48	0,61	0,65	0,64	0,61	0,58	0,57
"	0.2	0,51	0,62	0,65	0,65	0,63	0,60	0,55
"	0.4	0,55	0,62	0,64	0,64	0,62	0,59	0,50
"	0.6	0,56	0,61	0,63	0,63	0,61	0,57	0,46
"	0.8	0,56	0,61	0,62	0,62	0,60	0,55	0,44
"	1.0	0,56	0,60	0,61	0,61	0,59	0,54	0,42
"	2.0	0,56	0,59	0,59	0,58	0,56	0,50	0,37
"	4.0	0,55	0,57	0,57	0,57	0,54	0,47	0,34
"	6.0	0,54	0,57	0,57	0,56	0,53	0,46	0,33
"	8.0	0,54	0,56	0,56	0,55	0,52	0,46	0,32
Accessibility	-	0,55	0,56	0,55	0,54	0,51	0,44	0,30
Eigenvector	-	0,55	0,55	0,55	0,54	0,51	0,44	0,30

0,11 — 0,65

B) Multi-regional: South Sweden

Model version	α	β = 3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	0,10	0,17	0,27	0,46	0,48	0,47	0,35
"	0.002	0,10	0,17	0,27	0,48	0,49	0,47	0,36
"	0.004	0,11	0,18	0,29	0,51	0,50	0,47	0,36
"	0.006	0,11	0,19	0,32	0,53	0,51	0,47	0,36
"	0.008	0,11	0,19	0,36	0,53	0,51	0,48	0,36
"	0.01	0,11	0,19	0,38	0,54	0,52	0,48	0,36
"	0.02	0,12	0,20	0,44	0,55	0,54	0,47	0,37
"	0.04	0,12	0,19	0,47	0,56	0,55	0,48	0,37
"	0.06	0,12	0,19	0,49	0,57	0,55	0,48	0,37
"	0.08	0,12	0,19	0,48	0,57	0,56	0,48	0,36
"	0.1	0,12	0,19	0,49	0,57	0,55	0,48	0,35
"	0.2	0,12	0,36	0,48	0,56	0,55	0,48	0,29
"	0.4	0,27	0,35	0,53	0,56	0,54	0,44	0,22
"	0.6	0,26	0,36	0,53	0,56	0,52	0,41	0,18
"	0.8	0,27	0,44	0,53	0,55	0,51	0,38	0,16
"	1.0	0,27	0,42	0,52	0,54	0,50	0,36	0,15
"	2.0	0,27	0,41	0,51	0,52	0,46	0,31	0,12
"	4.0	0,31	0,41	0,50	0,50	0,43	0,28	0,10
"	6.0	0,30	0,41	0,50	0,50	0,42	0,26	0,10
"	8.0	0,30	0,41	0,50	0,49	0,42	0,26	0,09
Accessibility	-	0,46	0,50	0,52	0,49	0,40	0,24	0,09
Eigenvector	-	0,45	0,50	0,52	0,49	0,39	0,23	0,08

0,08 — 0,57

4. Discussion and conclusions

We have investigated how well a preferential centrality model can capture patterns of urban activity on the regional and multi-regional scales. This comparatively simple model incorporates important urban processes, such as the interplay between agglomeration and accessibility, which normally requires much more complicated models, for example LUTI- or spatial computable general equilibrium (SCGE) models.

Output from several model versions was compared to empirical data, both for a single region (Östergötland in Sweden) and for a larger multi-regional area (southern half of Sweden). Under the used performance indicators, the preferential model outperforms the simpler accessibility and eigenvector models, suggesting that while spatial interaction plays a crucial role, it is an irreducible interplay between spatial interaction and agglomeration that is in play, and the preferential centrality measure shows promise as a general basic model to capture these processes.

The results also demonstrate that the parameter β , controlling distance decay of interactions, has a major influence on how well the model fits empirical data. Values of β that represent substantial influence of long-range interactions produce the best empirical fit, confirming that urban land value variation should not be studied only as local phenomenon.

The results show strong correlations between road network structure and land taxation values and suggest that models without agglomeration feedback will tend to systematically underestimate these effects, especially for locations with high land values, such as regional and urban centres. The preferential model demonstrates that agglomeration and accessibility can be studied in an integrated fashion that captures interactions on many scales, using only a few key parameters to reproduce empirically observed spatial patterns

When comparing single regional- with multi-regional results, it is clear that model outputs fit empirical data better for the single region. This suggests that modelling on the multi-regional scale might require additional model mechanisms to take care of regional heterogeneity. One possibility would be to introduce regional variation of parameters reflecting different underlying economic structures. Other important factors that we have not studied, are the impacts of railway networks, air travel, and information technology. These could possibly have larger influence on the multi-regional scale compared to within a single region.

The travel time calculations used in this study are simplified, representing free flow car traffic moving according to speed limits. For a more realistic representation of urban transportation other modes such as walking, cycling, public transport as well as traffic congestion must also be considered. The choice to not include these aspects in the current study was based on the ambition to retain transparency and simplicity in the model structure. However, in future practical applications, it is clearly possible to use a separate full transportation model to generate realistic travel times for all relevant transport modes.

This study used land taxation data as a proxy for urban activity on a detailed spatial scale. Compared to other possible data sources, such as population or employment, the main advantage is that land taxation is not limited by physical constraints. This means that the differences between urban economic activity at different locations can be observed clearly. However, the usefulness of land taxation data rests on the assumption that taxation values give a reasonable representation of underlying land values. For many countries in the world this might not be the case, which points towards a need for further studies where preferential centrality is compared to other data sources, for example satellite night light data.

The present results are derived from a cross-sectional analysis with a static road network and a snapshot of land taxation data from one year. This means that the empirical observations cannot prove any causal link between network centrality and land value. Reverse causality or confounding factors, such as unobserved processes driving both network growth and land value, cannot be ruled out. However, since transport network growth takes place on a much slower timescale than land value change, it is plausible that network structure is an important underlying causal factor. To test this empirically, longitudinal studies are needed, where changes in the transport network are compared with changes in urban activity.

Another related future line of work is to study how the parameters for interaction (β) and agglomeration (α) can be estimated for specific model settings. For interaction, it is probably possible to use direct methods such as analysis of mobility data. Agglomeration, however, is a process that is more difficult to observe directly, raising the need to use indirect methods, such as the longitudinal studies suggested above. In these estimations it must also be considered that the forces of interaction and agglomeration might change over time due to changing technology and policy.

Although the aforementioned limitations of the reported empirical results, the preferential centrality model has many potential merits. The model can be used both as a framework to understand principal linkages between transport networks, accessibility and agglomeration, as well as a tool in practical planning processes. A possible application is the analysis of potential changes in urban activity due to alterations of the transport network. For example, preferential centrality calculations for a network with and without a proposed new road can provide insights into the geographical distribution of the economic impacts.

Compared to more complicated models, such as LUTI or SCGE, the lightweight approach of a preferential centrality model substantially lowers the barriers of entry into urban and regional modelling, especially in settings where data and resources are limited. This is achieved while retaining the important capacity to describe urban agglomeration processes, which is not present in other network centrality measures and static accessibility indicators.

5. References

Acheampong RA and Silva EA. (2015) Land use–transport interaction modeling: A review of the literature and future research directions. *Journal of Transport and Land Use* 8: 11-38.

Agryzkov T, Tortosa L, Vicent JF, et al. (2019) A centrality measure for urban networks based on the eigenvector centrality concept. *Environment and Planning B: Urban Analytics and City Science* 46: 668-689.

Alonso W. (1964) *Location and land use*: Harvard University Press.

Andersson C, Frenken K and Hellervik A. (2006) A complex network approach to urban growth. *Environment and Planning A* 38.

Barthelemy M. (2016) *The structure and dynamics of cities*: Cambridge University Press.

Barthélemy M. (2011) Spatial networks. *Physics Reports* 499: 1-101.

Baum-Snow N. (2007) Did highways cause suburbanization? *The quarterly journal of economics* 122: 775-805.

Boeing G. (2017) OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems* 65: 126-139.

Bonacich P. (1972) Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology* 2: 113-120.

De Montis A, Caschili S and Chessa A. (2013) Recent Developments of Complex Network Analysis in Spatial Planning. In: Scherngell T (ed) *The Geography of Networks and R&D Collaborations.* Cham: Springer International Publishing, 29-47.

Duranton G and Puga D. (2004) Micro-foundations of urban agglomeration economies. *Handbook of regional and urban economics.* Elsevier, 2063-2117.

Fujita M, Krugman PR, Venables AJ, et al. (1999) *The spatial economy: cities, regions and international trade*: Wiley Online Library.

Gibbons S, Lyytikäinen T, Overman HG, et al. (2019) New road infrastructure: the effects on firms. *Journal of Urban Economics* 110: 35-50.

Glaeser EL. (2008) *Cities, agglomeration, and spatial equilibrium*: Oxford University Press.

Handy SL and Niemeier DA. (1997) Measuring accessibility: an exploration of issues and alternatives. *Environment and Planning A* 29: 1175-1194.

Hansen WG. (1959) How Accessibility Shapes Land Use. *Journal of the American Institute of Planners* 25: 73-76.

Hellervik A, Nilsson L and Andersson C. (2019) Preferential centrality - a new measure unifying urban activity, attraction and accessibility. *Accepted for publication in Environment and Planning B: Urban Analytics and City Science.*

Hillier B and Hanson J. (1989) *The social logic of space*: Cambridge university press.

Jiang B. (2006) Ranking Spaces for Predicting Human Movement in an Urban Environment. *International Journal of Geographical Information Science* 23: 823-837.

Krugman P. (1996) Urban concentration: The role of increasing returns and transport costs. *International Regional Science Review* 19: 5-30.

Nelson KP, Kon MA and Umarov SR. (2019) Use of the geometric mean as a statistic for the scale of the coupled Gaussian distributions. *Physica A: Statistical Mechanics and its Applications* 515: 248-257.

Porta S, Crucitti P and Latora V. (2006) The network analysis of urban streets: a primal approach. *Environment and Planning B: Planning and Design* 33: 705-725.

Rey SJ and Anselin L. (2007) PySAL: A Python Library of Spatial Analytical Methods. *Review of Regional Studies* 37: 5-27.

Sevtsuk A and Mekonnen M. (2012) Urban network analysis. *Revue internationale de géomatique–n* 287: 305.

Statistics Sweden. (2016) SCB:s avgränsningar av koncentrerad bebyggelse.

Wegener M. (2004) Overview of land-use transport models. *Handbook of transport geography and spatial systems* 5: 127-146.

Preferential centrality as a multi-regional model for spatial interaction and urban agglomeration - supplementary regression results

This document contains a number of supplementary tables to provide additional information about the regression results. Results are presented for two separate modelling set-ups: A) The single Swedish region of Östergötland (N=185 observations), and B) a multi-regional area covering most of south Sweden (N=2831 observations).
The columns in the tables represent variation of the parameter β in the deterrence function, and the rows represent variation of model version and the parameter α.

All regression analyses were performed in PySAL (v 2.0.0) using ordinary least squares with spatial diagnostics. Nearest neighbor weights matrices were used, based on 4 nearest neighbors. The weights matrices were symmetrized and row-standardized.

All results refer to this regression model, described in the main article:

$$\log y_j^* = b_0 + b_1 \log a_j^* + \epsilon_j$$

Table 1, R^2-values for the OLS regression. (This table appears also in the main article.)

		A) Single region: Östergötland							B) Multi-regional: South Sweden						
Model version	α	$\beta =$ 3.5	3	2.5	2	1.5	1	0.5	$\beta =$ 3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	0.11	0.37	0.54	0.43	0.40	0.38	0.37	0.10	0.17	0.27	0.46	0.48	0.47	0.35
“	0.002	0.16	0.41	0.56	0.46	0.43	0.40	0.39	0.10	0.17	0.27	0.48	0.49	0.47	0.36
“	0.004	0.23	0.47	0.59	0.49	0.46	0.43	0.41	0.11	0.18	0.29	0.51	0.50	0.47	0.36
“	0.006	0.29	0.51	0.60	0.51	0.48	0.45	0.43	0.11	0.19	0.32	0.53	0.51	0.47	0.36
“	0.008	0.33	0.54	0.61	0.53	0.49	0.46	0.44	0.11	0.19	0.36	0.53	0.51	0.48	0.36
“	0.01	0.35	0.55	0.61	0.54	0.50	0.47	0.45	0.11	0.19	0.38	0.54	0.52	0.48	0.36
“	0.02	0.43	0.59	0.63	0.57	0.52	0.49	0.48	0.12	0.20	0.44	0.55	0.54	0.47	0.37
“	0.04	0.46	0.60	0.64	0.61	0.55	0.52	0.52	0.12	0.19	0.47	0.56	0.55	0.48	0.37
“	0.06	0.47	0.61	0.65	0.63	0.58	0.54	0.55	0.12	0.19	0.49	0.57	0.55	0.48	0.37
“	0.08	0.48	0.61	0.65	0.64	0.60	0.56	0.56	0.12	0.19	0.48	0.57	0.56	0.48	0.36
“	0.1	0.48	0.61	0.65	0.64	0.61	0.58	0.57	0.12	0.19	0.49	0.57	0.55	0.48	0.35
“	0.2	0.51	0.62	0.65	0.65	0.63	0.60	0.55	0.12	0.36	0.48	0.56	0.55	0.48	0.29
“	0.4	0.55	0.62	0.64	0.64	0.62	0.59	0.50	0.27	0.35	0.53	0.56	0.54	0.44	0.22
“	0.6	0.56	0.61	0.63	0.63	0.61	0.57	0.46	0.26	0.36	0.53	0.56	0.52	0.41	0.18
“	0.8	0.56	0.61	0.62	0.62	0.60	0.55	0.44	0.27	0.44	0.53	0.55	0.51	0.38	0.16
“	1.0	0.56	0.60	0.61	0.61	0.59	0.54	0.42	0.27	0.42	0.52	0.54	0.50	0.36	0.15
“	2.0	0.56	0.59	0.59	0.58	0.56	0.50	0.37	0.27	0.41	0.51	0.52	0.46	0.31	0.12
“	4.0	0.55	0.57	0.57	0.57	0.54	0.47	0.34	0.31	0.41	0.50	0.50	0.43	0.28	0.10
“	6.0	0.54	0.57	0.57	0.56	0.53	0.46	0.33	0.30	0.41	0.50	0.50	0.42	0.26	0.10
“	8.0	0.54	0.56	0.56	0.55	0.52	0.46	0.32	0.30	0.41	0.50	0.49	0.42	0.26	0.09
Accessibility	-	0.55	0.56	0.55	0.54	0.51	0.44	0.30	0.46	0.50	0.52	0.49	0.40	0.24	0.09
Eigenvector	-	0.55	0.55	0.55	0.54	0.51	0.44	0.30	0.45	0.50	0.52	0.49	0.39	0.23	0.08

0.11 0.65

0.08 0.57

Table 2, OLS estimation for the slope coefficient b_1

		A) Single region: Östergötland							B) Multi-regional: South Sweden						
		$\beta =$							$\beta =$						
Model version	α	3.5	3	2.5	2	1.5	1	0.5	3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	0.18	0.32	0.43	0.40	0.42	0.43	0.40	0.60	0.62	0.55	0.45	0.39	0.37	0.28
"	0.002	0.24	0.37	0.49	0.45	0.47	0.48	0.45	0.62	0.65	0.57	0.52	0.45	0.42	0.32
"	0.004	0.33	0.45	0.54	0.51	0.52	0.53	0.51	0.66	0.70	0.43	0.60	0.51	0.48	0.37
"	0.006	0.39	0.49	0.56	0.54	0.55	0.56	0.54	0.68	0.73	0.49	0.64	0.56	0.53	0.41
"	0.008	0.43	0.52	0.58	0.56	0.57	0.57	0.56	0.71	0.76	0.56	0.67	0.59	0.55	0.44
"	0.01	0.46	0.54	0.59	0.57	0.58	0.59	0.58	0.72	0.79	0.60	0.69	0.62	0.58	0.48
"	0.02	0.54	0.59	0.61	0.61	0.61	0.62	0.65	0.78	0.87	0.73	0.77	0.72	0.67	0.60
"	0.04	0.57	0.61	0.63	0.64	0.64	0.67	0.77	0.82	0.98	0.80	0.82	0.82	0.78	0.79
"	0.06	0.59	0.62	0.64	0.66	0.68	0.72	0.88	0.84	1.01	0.84	0.85	0.89	0.87	0.94
"	0.08	0.60	0.62	0.64	0.68	0.70	0.76	0.97	0.85	1.03	0.85	0.87	0.93	0.94	1.06
"	0.1	0.60	0.63	0.65	0.69	0.73	0.81	1.05	0.86	1.05	0.87	0.89	0.97	1.00	1.17
"	0.2	0.64	0.65	0.68	0.73	0.82	0.96	1.30	0.88	0.76	0.92	0.97	1.07	1.21	1.43
"	0.4	0.68	0.69	0.72	0.80	0.92	1.14	1.57	0.67	0.81	1.00	1.04	1.20	1.45	1.54
"	0.6	0.71	0.71	0.76	0.85	1.00	1.25	1.69	0.70	0.85	1.04	1.10	1.29	1.56	1.53
"	0.8	0.73	0.74	0.78	0.88	1.05	1.33	1.76	0.72	0.93	1.08	1.15	1.35	1.63	1.51
"	1.0	0.74	0.76	0.80	0.91	1.09	1.38	1.80	0.74	0.94	1.11	1.19	1.40	1.67	1.48
"	2.0	0.79	0.81	0.87	0.99	1.20	1.52	1.87	0.79	1.00	1.17	1.29	1.52	1.73	1.40
"	4.0	0.83	0.86	0.92	1.06	1.28	1.60	1.89	0.84	1.05	1.22	1.37	1.60	1.75	1.34
"	6.0	0.85	0.87	0.94	1.08	1.31	1.63	1.90	0.86	1.06	1.24	1.40	1.63	1.75	1.31
"	8.0	0.86	0.88	0.95	1.09	1.33	1.65	1.90	0.86	1.07	1.25	1.41	1.65	1.74	1.30
Accessibility	-	0.83	0.89	0.98	1.14	1.38	1.70	1.89	1.00	1.10	1.23	1.43	1.69	1.74	1.25
Eigenvector	-	0.86	0.90	0.98	1.14	1.38	1.70	1.89	0.99	1.10	1.23	1.44	1.69	1.73	1.25

0.18 1.90
1.00

0.28 1.75
1.00

Table 3, p-values for the slope coefficient b_1

		A) Single region: Östergötland							B) Multi-regional: South Sweden						
		$\beta =$							$\beta =$						
Model version	α	3.5	3	2.5	2	1.5	1	0.5	3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	4.3E-06	2.5E-20	1.9E-32	4.3E-24	3.6E-22	1.0E-20	8.3E-20	6.2E-68	1.8E-117	5.3E-193	0.0E+00	0.0E+00	0.0E+00	2.7E-271
"	0.002	2.3E-08	9.7E-23	7.9E-35	3.4E-26	5.7E-24	2.9E-22	3.6E-21	1.7E-69	2.6E-120	5.4E-197	0.0E+00	0.0E+00	0.0E+00	6.9E-273
"	0.004	3.5E-12	3.3E-27	3.7E-37	1.1E-28	4.8E-26	4.4E-24	7.8E-23	7.8E-72	9.3E-125	6.8E-211	0.0E+00	0.0E+00	0.0E+00	7.4E-275
"	0.006	4.2E-15	2.3E-30	2.9E-38	2.6E-30	1.7E-27	2.7E-25	6.3E-24	1.6E-73	4.5E-128	7.3E-238	0.0E+00	0.0E+00	0.0E+00	1.7E-276
"	0.008	2.4E-17	1.8E-32	5.3E-39	1.5E-31	2.1E-28	4.6E-26	9.2E-25	8.7E-75	1.2E-130	5.9E-273	0.0E+00	0.0E+00	0.0E+00	6.0E-278
"	0.01	4.2E-19	5.9E-34	1.6E-39	1.5E-32	4.2E-29	9.8E-27	2.0E-25	8.8E-76	1.0E-132	2.2E-291	0.0E+00	0.0E+00	0.0E+00	1.4E-278
"	0.02	6.9E-24	2.5E-37	3.7E-41	9.9E-36	3.0E-31	1.2E-28	1.3E-27	1.3E-78	7.2E-139	0.0E+00	0.0E+00	0.0E+00	0.0E+00	3.1E-282
"	0.04	2.1E-26	1.7E-38	1.3E-42	2.8E-39	6.1E-34	5.4E-31	8.2E-31	9.9E-81	2.0E-131	0.0E+00	0.0E+00	0.0E+00	0.0E+00	6.6E-285
"	0.06	3.8E-27	2.9E-39	3.1E-43	4.0E-41	2.8E-36	4.5E-33	3.6E-33	1.4E-81	9.4E-133	0.0E+00	0.0E+00	0.0E+00	0.0E+00	9.1E-284
"	0.08	5.4E-28	2.3E-39	1.5E-43	5.1E-42	5.9E-38	7.6E-35	1.9E-34	5.1E-82	1.8E-133	0.0E+00	0.0E+00	0.0E+00	0.0E+00	6.0E-279
"	0.1	5.1E-28	1.3E-39	1.0E-43	1.4E-42	4.7E-39	3.7E-36	5.1E-35	2.7E-82	6.1E-134	0.0E+00	0.0E+00	0.0E+00	0.0E+00	2.0E-271
"	0.2	4.3E-30	5.6E-40	1.0E-43	2.0E-43	3.8E-41	2.3E-38	1.2E-33	7.4E-83	3.6E-277	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.3E-216
"	0.4	2.6E-33	4.1E-40	1.3E-42	1.1E-42	9.5E-41	5.5E-37	3.2E-29	5.3E-199	1.6E-270	0.0E+00	0.0E+00	0.0E+00	0.0E+00	6.4E-154
"	0.6	2.6E-34	1.3E-39	2.0E-41	2.2E-41	2.3E-39	4.7E-35	2.3E-26	2.5E-189	6.0E-273	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.1E-125
"	0.8	1.7E-34	6.2E-39	2.1E-40	2.9E-40	3.9E-38	1.5E-33	1.7E-24	4.2E-198	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.7E-298	3.7E-110
"	1.0	1.4E-34	1.9E-38	1.2E-39	2.4E-39	3.7E-37	1.9E-32	3.7E-23	1.6E-198	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.0E-278	7.4E-100
"	2.0	5.1E-34	8.6E-37	2.2E-37	8.7E-37	2.1E-34	2.0E-29	5.2E-20	3.5E-198	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.1E-230	1.1E-78
"	4.0	2.8E-33	1.7E-35	8.8E-36	5.6E-35	1.7E-32	2.2E-27	4.1E-18	7.7E-227	0.0E+00	0.0E+00	0.0E+00	0.0E+00	6.6E-201	7.6E-68
"	6.0	6.9E-33	5.8E-35	3.9E-35	2.9E-34	9.4E-32	1.3E-26	2.0E-17	8.9E-221	0.0E+00	0.0E+00	0.0E+00	0.0E+00	7.5E-190	5.4E-64
"	8.0	9.8E-33	1.1E-34	8.2E-35	6.7E-34	2.3E-31	3.4E-26	4.4E-17	1.3E-217	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.7E-184	3.2E-62
Accessibility	-	3.1E-33	5.0E-34	6.1E-34	7.6E-33	3.6E-30	6.6E-25	5.0E-16	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.8E-168	5.2E-57
Eigenvector	-	3.2E-33	5.8E-34	7.8E-34	9.0E-33	4.1E-30	7.5E-25	6.0E-16	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	2.4E-165	2.8E-56

Table 4, Spatial diagnostics - Robust lagrange multiplier test for spatial lag model, statistic

		A) Single region: Östergötland							B) Multi-regional: South Sweden						
		$\beta =$							$\beta =$						
Model version	α	3.5	3	2.5	2	1.5	1	0.5	3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	0.49	6.37	11.60	4.91	5.44	5.53	3.49	42.41	46.05	27.32	2.94	0.35	3.00	22.44
“	0.002	1.76	6.25	11.56	4.88	5.02	4.75	2.82	43.57	48.09	28.43	3.85	0.63	1.47	17.31
“	0.004	4.77	7.72	11.40	4.65	4.21	3.68	1.94	45.65	51.96	0.00	5.08	1.21	0.44	11.98
“	0.006	7.54	9.37	10.53	4.39	3.70	3.01	1.44	47.42	55.47	0.35	5.81	1.84	0.10	8.78
“	0.008	9.87	10.52	9.91	4.13	3.20	2.48	1.14	48.92	58.61	1.90	6.40	2.47	0.03	6.50
“	0.01	11.78	11.24	9.38	3.93	2.81	2.13	0.94	50.22	61.43	3.37	6.96	3.06	0.01	4.41
“	0.02	16.92	12.30	7.87	3.38	1.77	1.18	0.60	54.70	71.96	13.57	10.13	6.79	1.05	0.25
“	0.04	19.26	12.30	6.88	3.15	1.24	0.77	1.06	59.11	90.11	25.20	14.36	12.80	5.86	3.75
“	0.06	19.97	12.38	6.58	3.18	1.32	0.92	2.46	61.31	95.58	32.79	17.64	19.53	12.67	15.52
“	0.08	20.73	12.58	6.51	3.26	1.57	1.35	4.42	62.64	99.17	36.20	20.67	24.33	20.60	32.10
“	0.1	21.00	12.72	6.53	3.40	1.88	1.94	6.57	63.54	101.75	41.25	23.71	29.62	29.21	50.98
“	0.2	23.05	13.49	7.10	4.40	3.68	5.61	16.44	65.62	42.65	56.91	39.09	48.46	74.45	123.54
“	0.4	24.67	14.33	8.61	6.56	7.11	12.04	29.48	39.83	57.85	78.48	54.34	80.26	147.61	175.96
“	0.6	25.45	15.32	9.95	8.33	9.76	16.72	36.81	48.97	70.30	93.71	68.50	106.39	192.29	189.45
“	0.8	26.20	16.45	11.06	9.70	11.84	20.17	41.21	53.05	102.10	106.84	83.14	127.13	219.87	190.95
“	1.0	26.74	17.21	11.95	10.80	13.46	22.78	44.05	57.58	105.37	116.35	95.31	143.57	237.82	187.79
“	2.0	28.83	19.85	14.67	14.01	18.15	29.70	49.78	70.59	123.87	141.46	128.67	190.25	275.63	169.51
“	4.0	30.59	22.01	16.83	16.50	21.61	34.29	52.33	91.64	144.18	159.87	154.66	223.46	291.28	152.14
“	6.0	31.44	22.95	17.77	17.57	23.03	36.04	53.06	93.62	152.17	166.67	166.75	235.95	294.84	144.36
“	8.0	31.77	23.32	18.20	18.11	23.77	36.92	53.37	94.73	156.41	170.31	172.09	242.86	296.35	140.53
Accessibility	-	25.86	21.66	18.95	19.74	26.14	39.69	54.15	207.24	180.56	157.79	178.60	263.49	299.28	128.30
Eigenvector	-	28.63	22.59	19.18	19.82	26.24	39.79	54.19	197.94	176.80	156.41	179.71	265.54	297.30	126.31

0.5 — 54.2

0.0 — 299.3

Table 5, Spatial diagnostics - Robust lagrange multiplier test for spatial lag model, p-value

A) Single region: Östergötland

Model version	α	$\beta =$ 3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	0.48607	0.01164	0.00066	0.02676	0.01970	0.01872	0.06181
"	0.002	0.18465	0.01245	0.00067	0.02710	0.02499	0.02933	0.09319
"	0.004	0.02888	0.00545	0.00073	0.03099	0.04008	0.05517	0.16409
"	0.006	0.00605	0.00221	0.00117	0.03616	0.05442	0.08250	0.23060
"	0.008	0.00168	0.00118	0.00164	0.04219	0.07370	0.11532	0.28480
"	0.01	0.00060	0.00080	0.00219	0.04751	0.09389	0.14484	0.33116
"	0.02	0.00004	0.00045	0.00502	0.06618	0.18374	0.27837	0.43797
"	0.04	0.00001	0.00045	0.00872	0.07574	0.26556	0.38052	0.30391
"	0.06	0.00001	0.00043	0.01030	0.07463	0.25018	0.33693	0.11650
"	0.08	0.00001	0.00039	0.01074	0.07108	0.20986	0.24574	0.03556
"	0.1	0.00000	0.00036	0.01058	0.06521	0.17065	0.16417	0.01037
"	0.2	0.00000	0.00024	0.00772	0.03593	0.05520	0.01785	0.00005
"	0.4	0.00000	0.00015	0.00335	0.01042	0.00768	0.00052	0.00000
"	0.6	0.00000	0.00009	0.00161	0.00391	0.00178	0.00004	0.00000
"	0.8	0.00000	0.00005	0.00088	0.00185	0.00058	0.00001	0.00000
"	1.0	0.00000	0.00003	0.00055	0.00102	0.00024	0.00000	0.00000
"	2.0	0.00000	0.00001	0.00013	0.00018	0.00002	0.00000	0.00000
"	4.0	0.00000	0.00000	0.00004	0.00005	0.00000	0.00000	0.00000
"	6.0	0.00000	0.00000	0.00002	0.00003	0.00000	0.00000	0.00000
"	8.0	0.00000	0.00000	0.00002	0.00002	0.00000	0.00000	0.00000
Accessibility	-	0.00000	0.00000	0.00001	0.00001	0.00000	0.00000	0.00000
Eigenvector	-	0.00000	0.00000	0.00001	0.00001	0.00000	0.00000	0.00000

B) Multi-regional: South Sweden

Model version	α	$\beta =$ 3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	0.00000	0.00000	0.00000	0.08637	0.55373	0.08344	0.00000
"	0.002	0.00000	0.00000	0.00000	0.04971	0.42793	0.22504	0.00003
"	0.004	0.00000	0.00000	0.98085	0.02426	0.27097	0.50691	0.00054
"	0.006	0.00000	0.00000	0.55440	0.01592	0.17468	0.75263	0.00304
"	0.008	0.00000	0.00000	0.16753	0.01140	0.11620	0.86749	0.01079
"	0.01	0.00000	0.00000	0.06655	0.00833	0.08033	0.93337	0.03574
"	0.02	0.00000	0.00000	0.00023	0.00146	0.00915	0.30505	0.61708
"	0.04	0.00000	0.00000	0.00000	0.00015	0.00035	0.01547	0.05271
"	0.06	0.00000	0.00000	0.00000	0.00003	0.00001	0.00037	0.00008
"	0.08	0.00000	0.00000	0.00000	0.00001	0.00000	0.00001	0.00000
"	0.1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
"	0.2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
"	0.4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
"	0.6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
"	0.8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
"	1.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
"	2.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
"	4.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
"	6.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
"	8.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Accessibility	-	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Eigenvector	-	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

0.00 — 0.98

0.00 — 0.98

Table 6, Spatial diagnostics - Robust lagrange multiplier test for spatial error model, statistic

		A) Single region: Östergötland							B) Multi-regional: South Sweden						
		$\beta =$							$\beta =$						
Model version	α	3.5	3	2.5	2	1.5	1	0.5	3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	9.69	8.08	1.01	4.70	4.28	3.77	3.57	143.01	248.80	346.56	234.74	163.29	176.39	176.24
“	0.002	9.57	8.17	0.48	4.50	4.49	4.33	4.37	146.22	252.76	350.40	255.39	183.89	190.02	182.05
“	0.004	8.44	5.77	0.13	4.23	4.84	5.03	5.37	150.57	257.81	348.82	274.43	208.71	209.71	193.26
“	0.006	7.03	3.53	0.12	4.06	5.02	5.48	6.02	153.29	260.67	367.82	284.53	223.75	223.33	203.50
“	0.008	5.72	2.18	0.11	3.95	5.29	5.88	6.48	155.08	262.30	385.98	291.22	233.96	231.98	212.86
“	0.01	4.66	1.42	0.11	3.82	5.52	6.16	6.82	156.31	263.19	394.02	296.04	241.38	240.10	219.87
“	0.02	2.05	0.33	0.14	3.26	6.15	7.01	7.49	158.81	263.18	405.83	308.46	264.57	266.25	253.84
“	0.04	1.08	0.19	0.14	2.14	5.92	7.13	6.19	159.26	231.95	398.06	313.84	280.81	291.31	299.33
“	0.06	0.84	0.09	0.12	1.48	4.80	6.16	3.76	158.90	228.70	390.44	315.69	287.05	305.35	324.35
“	0.08	0.59	0.07	0.10	1.15	3.73	4.81	1.93	158.51	226.29	389.58	316.85	288.89	311.93	334.03
“	0.1	0.55	0.04	0.08	0.90	2.92	3.57	0.85	158.18	224.43	384.55	317.81	290.23	314.77	333.49
“	0.2	0.10	0.00	0.03	0.35	0.90	0.61	0.13	157.22	398.66	375.51	319.15	290.01	306.89	279.95
“	0.4	0.08	0.03	0.01	0.08	0.10	0.01	1.67	365.49	384.42	329.80	309.85	283.61	274.37	184.98
“	0.6	0.28	0.05	0.01	0.04	0.01	0.22	2.87	347.91	374.39	318.08	303.09	276.73	249.79	133.40
“	0.8	0.43	0.08	0.01	0.02	0.00	0.49	3.65	351.19	350.75	313.04	298.68	270.89	231.14	105.25
“	1.0	0.55	0.10	0.01	0.01	0.02	0.75	4.17	346.95	349.59	306.70	295.04	265.95	216.58	88.13
“	2.0	0.84	0.18	0.00	0.00	0.17	1.57	5.30	333.73	334.54	291.67	283.35	248.31	175.10	57.51
“	4.0	1.05	0.26	0.00	0.01	0.36	2.23	5.90	332.34	317.32	282.40	273.30	232.19	146.32	45.42
“	6.0	1.14	0.31	0.01	0.02	0.46	2.49	6.10	326.98	310.58	278.64	268.31	225.04	135.29	41.88
“	8.0	1.19	0.32	0.01	0.03	0.51	2.63	6.20	323.97	307.01	276.64	265.95	220.96	129.70	40.35
Accessibility	-	0.36	0.10	0.02	0.07	0.72	3.09	6.48	256.85	267.08	274.30	260.63	207.29	112.48	36.37
Eigenvector	-	0.78	0.17	0.02	0.07	0.73	3.10	6.50	266.13	271.93	277.11	261.94	205.94	110.41	35.97

0.00 9.69

35.97 405.83

Table 7, Spatial diagnostics - Robust lagrange multiplier test for spatial error model, p-value

A) Single region: Östergötland

Model version	α	$\beta =$ 3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	0.00186	0.00448	0.31601	0.03016	0.03864	0.05227	0.05872
"	0.002	0.00198	0.00426	0.48878	0.03393	0.03416	0.03742	0.03648
"	0.004	0.00367	0.01628	0.71940	0.03972	0.02785	0.02498	0.02045
"	0.006	0.00804	0.06044	0.73193	0.04387	0.02501	0.01919	0.01412
"	0.008	0.01674	0.13984	0.74310	0.04689	0.02147	0.01536	0.01090
"	0.01	0.03080	0.23381	0.73764	0.05057	0.01883	0.01308	0.00901
"	0.02	0.15208	0.56793	0.70625	0.07102	0.01318	0.00811	0.00621
"	0.04	0.29875	0.66682	0.70784	0.14328	0.01493	0.00760	0.01282
"	0.06	0.35820	0.76512	0.72817	0.22407	0.02842	0.01307	0.05248
"	0.08	0.44135	0.79053	0.74731	0.28457	0.05336	0.02826	0.16510
"	0.1	0.45647	0.83924	0.77083	0.34193	0.08758	0.05897	0.35756
"	0.2	0.74985	0.98624	0.85384	0.55619	0.34274	0.43491	0.71987
"	0.4	0.78258	0.87129	0.90891	0.77203	0.75231	0.92557	0.19589
"	0.6	0.59898	0.81501	0.91973	0.84411	0.93666	0.63881	0.09018
"	0.8	0.51370	0.77591	0.92449	0.88089	0.95416	0.48278	0.05617
"	1.0	0.45977	0.75210	0.93443	0.90469	0.87645	0.38730	0.04122
"	2.0	0.35905	0.67527	0.98459	0.99126	0.67994	0.20962	0.02130
"	4.0	0.30521	0.61040	0.95369	0.92539	0.54870	0.13561	0.01516
"	6.0	0.28506	0.58004	0.92101	0.88598	0.49888	0.11429	0.01353
"	8.0	0.27580	0.57153	0.90697	0.86460	0.47331	0.10454	0.01281
Accessibility	-	0.55030	0.74972	0.90115	0.79460	0.39490	0.07864	0.01090
Eigenvector	-	0.37625	0.67838	0.89185	0.79562	0.39380	0.07806	0.01079

0.00 – 0.99

B) Multi-regional: South Sweden

Model version	α	$\beta =$ 3.5	3	2.5	2	1.5	1	0.5
Preferential	0.001	5.86E-33	4.74E-56	2.38E-77	5.51E-53	2.16E-37	2.97E-40	3.21E-40
"	0.002	1.16E-33	6.49E-57	3.47E-78	1.74E-57	6.87E-42	3.15E-43	1.73E-41
"	0.004	1.30E-34	5.14E-58	7.66E-78	1.23E-61	2.62E-47	1.59E-47	6.18E-44
"	0.006	3.32E-35	1.23E-58	5.57E-82	7.73E-64	1.37E-50	1.70E-50	3.60E-46
"	0.008	1.34E-35	5.42E-59	6.20E-86	2.69E-65	8.17E-53	2.21E-52	3.26E-48
"	0.01	7.23E-36	3.46E-59	1.10E-87	2.41E-66	1.97E-54	3.75E-54	9.66E-50
"	0.02	2.06E-36	3.47E-59	2.97E-90	4.72E-69	1.73E-59	7.46E-60	3.78E-57
"	0.04	1.64E-36	2.24E-52	1.46E-88	3.19E-70	5.00E-63	2.58E-65	4.61E-67
"	0.06	1.97E-36	1.15E-51	6.65E-87	1.26E-70	2.18E-64	2.25E-68	1.64E-72
"	0.08	2.39E-36	3.85E-51	1.02E-86	7.04E-71	8.67E-65	8.30E-70	1.28E-74
"	0.1	2.82E-36	9.77E-51	1.27E-85	4.35E-71	4.43E-65	1.99E-70	1.67E-74
"	0.2	4.59E-36	1.08E-88	1.18E-83	2.22E-71	4.95E-65	1.04E-68	7.70E-63
"	0.4	1.80E-81	1.36E-85	1.06E-73	2.35E-69	1.22E-63	1.27E-61	3.96E-42
"	0.6	1.21E-77	2.07E-83	3.79E-71	7.01E-68	3.87E-62	2.89E-56	7.41E-31
"	0.8	2.33E-78	2.91E-78	4.76E-70	6.38E-67	7.27E-61	3.36E-52	1.08E-24
"	1.0	1.96E-77	5.21E-78	1.14E-68	3.96E-66	8.67E-60	5.04E-49	6.13E-21
"	2.0	1.48E-74	9.86E-75	2.15E-65	1.40E-63	6.06E-56	5.69E-40	3.37E-14
"	4.0	2.98E-74	5.55E-71	2.25E-63	2.17E-61	1.98E-52	1.10E-33	1.59E-11
"	6.0	4.37E-73	1.63E-69	1.49E-62	2.65E-60	7.19E-51	2.86E-31	9.69E-11
"	8.0	1.98E-72	9.79E-69	4.04E-62	8.68E-60	5.59E-50	4.77E-30	2.12E-10
Accessibility	-	8.35E-58	4.90E-60	1.31E-61	1.25E-58	5.35E-47	2.81E-26	1.63E-09
Eigenvector	-	7.92E-60	4.31E-61	3.20E-62	6.48E-59	1.05E-46	7.95E-26	2.01E-09

0.00 – 0.99